

47CD



OIPE

ENTERED

DATE: 07/05/2002 P.6  
TIME: 14:47:36RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/855,604AInput Set : D:\37156201.app  
Output Set: N:\CRF3\07052002\I855604A.raw

3 <110> APPLICANT: GICQUEL, BRIGITTE  
 4 PORTNOI, DENIS  
 5 LIM, ENG-MONG  
 6 PELICIC, VLADIMIR  
 7 GUIGUENO, AGNES  
 8 GOGUET DE LA SALMONIERE, YVES  
 10 <120> TITLE OF INVENTION: POLYPEPTIDE NUCLEIC SEQUENCES EXPORTED FROM MYCOBACTERIA,  
 11 VECTORS COMPRISING SAME AND USES FOR DIAGNOSING AND  
 12 PREVENTING TUBERCULOSIS  
 14 <130> FILE REFERENCE: 03715.0062-01000  
 16 <140> CURRENT APPLICATION NUMBER: 09/855,604A  
 17 <141> CURRENT FILING DATE: 2001-05-16  
 19 <150> PRIOR APPLICATION NUMBER: 09/485,536  
 20 <151> PRIOR FILING DATE: 2000-02-14  
 22 <150> PRIOR APPLICATION NUMBER: PCT/FR98/01813  
 23 <151> PRIOR FILING DATE: 1998-08-14  
 25 <150> PRIOR APPLICATION NUMBER: FR 97 10404  
 26 <151> PRIOR FILING DATE: 1997-08-14  
 28 <150> PRIOR APPLICATION NUMBER: FR 97 11325  
 29 <151> PRIOR FILING DATE: 1997-09-11  
 31 <160> NUMBER OF SEQ ID NOS: 935  
 33 <170> SOFTWARE: PatentIn Ver. 2.1  
 35 <210> SEQ ID NO: 1  
 36 <211> LENGTH: 1243  
 37 <212> TYPE: DNA  
 38 <213> ORGANISM: Mycobacterium tuberculosis  
 40 <220> FEATURE:  
 41 <221> NAME/KEY: CDS  
 42 <222> LOCATION: (21)..(38)  
 44 <220> FEATURE:  
 45 <221> NAME/KEY: CDS  
 46 <222> LOCATION: (124)..(453)  
 48 <220> FEATURE:  
 49 <221> NAME/KEY: CDS  
 50 <222> LOCATION: (473)..(481)  
 52 <220> FEATURE:  
 53 <221> NAME/KEY: CDS  
 54 <222> LOCATION: (573)..(671)  
 56 <220> FEATURE:  
 57 <221> NAME/KEY: CDS  
 58 <222> LOCATION: (719)..(928)  
 60 <220> FEATURE:  
 61 <221> NAME/KEY: CDS

RECEIVED

JUL 24 2002

OFFICE OF PETITIONS

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/855,604A

DATE: 07/05/2002  
TIME: 14:47:36

Input Set : D:\37156201.app  
Output Set: N:\CRF3\07052002\I855604A.raw

62 <222> LOCATION: (942)..(1241)  
 64 <400> SEQUENCE: 1  
 65 ggatccagg gaacgtgacc atg gtc gta ggg atg act tgacagttc 48  
 66 Met Val Val Gly Met Thr  
 67 1 5  
 69 aacggggtgc gaccaccgtt gcgcctcagaa ggcatacggtt ggtggaaacac gtcggaaagc 108  
 71 tggggagtga atctg atg gct ggc gac caa gag ctg gaa ctg cgg ttc gac 159  
 72 Met Ala Gly Asp Gln Glu Leu Glu Leu Arg Phe Asp  
 73 10 15  
 75 gtt cct ctt tac acg ctt gcc gag gca tcg cgg tac ctg gtg gtt ccc 207  
 76 Val Pro Leu Tyr Thr Leu Ala Glu Ala Ser Arg Tyr Leu Val Val Pro  
 77 20 25 30  
 79 cgc gcc acc ctg gct acg tgg gct gac ggc tac gag cgt cgg ccg gcc 255  
 80 Arg Ala Thr Leu Ala Thr Trp Ala Asp Gly Tyr Glu Arg Arg Pro Ala  
 81 35 40 45 50  
 83 aac gca ccg gcg gtc cag ggg caa ccg atc gcc ttt gac gcc tat tcg 303  
 84 Asn Ala Pro Ala Val Gln Gly Gln Pro Ile Ala Phe Asp Ala Tyr Ser  
 85 55 60 65  
 87 gtc gcg cag ctt ttt ggc gac gtc act ggt gcc cgc gtt gcg ggc gtc 351  
 88 Val Ala Gln Leu Phe Gly Asp Val Thr Gly Ala Arg Val Ala Gly Val  
 89 70 75 80  
 91 cag ccg cag cga cac cac ata cgg ccg gtc cgg ttg cgg ggg ccg ttg 399  
 92 Gln Pro Gln Arg His His Ile Arg Pro Val Arg Leu Arg Gly Pro Leu  
 93 85 90 95  
 95 ggt ggg gtt ggg tgc ctc cgt cac ccc agg cag ttc gct ggc tat ttg 447  
 96 Gly Gly Val Gly Cys Leu Arg His Pro Arg Gln Phe Ala Gly Tyr Leu  
 97 100 105 110  
 99 tcg cag tagcgcgacg gcattgtcg atg tct tgg tagctagcat ccggtcgggg 501  
 100 Ser Gln 115 Met Ser Trp  
 101 115  
 103 ggccgcgtacc agcgccagcg ccggggctcc ccgggtccggg tagtgcgcgt cgagttggtc 561  
 105 gtggaccagc a atg act gcg acc cgg cga ctt cga aac cgc cac cgg tta 611  
 106 Met Thr Ala Thr Arg Arg Leu Arg Asn Arg His Arg Leu  
 107 120 125 130  
 109 gat tcc ccg act gcg tca tcg cca ggt aaa ccg ccg gca cta acg cca 659  
 110 Asp Ser Pro Thr Ala Ser Ser Pro Gly Lys Pro Pro Ala Leu Thr Pro  
 111 135 140 145  
 113 gca acc aac ccg tgaagaccaa ccaacggcac ctgcgcaggt tgccggctcaa 711  
 114 Ala Thr Asn Pro  
 115 150  
 117 ccgcatac atg aac tgc tgg att tcg gac tcc ccg tac tct cgc gca gtg 760  
 118 Met Asn Cys Trp Ile Ser Asp Ser Pro Tyr Ser Arg Ala Val  
 119 155 160 165  
 121 cgt gcc ccg gag cct acc gaa gat cgc gtg cat gcg ttc ggc gtg gac 808  
 122 Arg Ala Arg Glu Pro Thr Glu Asp Arg Val His Ala Phe Gly Val Asp  
 123 170 175 180  
 125 cgc aca gca cct gga gtt ggc ggc gcc gag ggc cga gat ggc agg atg 856  
 126 Arg Thr Ala Pro Gly Val Gly Ala Glu Gly Arg Asp Gly Arg Met  
 127 185 190 195

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/855,604A

DATE: 07/05/2002  
TIME: 14:47:36

Input Set : D:\37156201.app  
Output Set: N:\CRF3\07052002\I855604A.raw

129 acg gat cgt cgg ggg cgg gaa ctc cca ggc cgc cgg acc gtc gca aac	904
130 Thr Asp Arg Arg Gly Arg Glu Leu Pro Gly Arg Arg Thr Val Ala Asn	
131 200 205 210	
133 ccc tcg caa acc cgt cgc aaa ccg taaggagtca tcc atg aag aca ggc	953
134 Pro Ser Gln Thr Arg Arg Lys Pro	Met Lys Thr Gly
135 215 220 225	
137 acc gcg acg acg cgg cgc agg ctg ttg gca gta ctg atc gcc ctc gcg	1001
138 Thr Ala Thr Thr Arg Arg Leu Leu Ala Val Leu Ile Ala Leu Ala	
139 230 235 240	
141 ttg ccg ggg gcc gcc gtt gcg ctg ctg gcc gaa cca tca gcg acc ggc	1049
142 Leu Pro Gly Ala Ala Val Ala Leu Leu Ala Glu Pro Ser Ala Thr Gly	
143 245 250 255	
145 gcg tcg gac ccg tgc gcg gcc agc gaa gtg gcg agg acg gtc ggt tcg	1097
146 Ala Ser Asp Pro Cys Ala Ala Ser Glu Val Ala Arg Thr Val Gly Ser	
147 260 265 270	
149 gtc gcc aag tcg atg ggc gac tac ctg gat tca cac cca gag acc aac	1145
150 Val Ala Lys Ser Met Gly Asp Tyr Leu Asp Ser His Pro Glu Thr Asn	
151 275 280 285 290	
153 cag gtg atg acc gcg gtc ttg cag cag cag gta ggg ccg ggg tcg gtc	1193
154 Gln Val Met Thr Ala Val Leu Gln Gln Val Gly Pro Gly Ser Val	
155 295 300 305	
157 gca tcg ctg aag gcc cat ttc gag gcg aat ccc aag gtc gca tcg gat	1241
158 Ala Ser Leu Lys Ala His Phe Glu Ala Asn Pro Lys Val Ala Ser Asp	
159 310 315 320	
161 cc	1243
164 <210> SEQ ID NO: 2	
165 <211> LENGTH: 322	
166 <212> TYPE: PRT	
167 <213> ORGANISM: Mycobacterium tuberculosis	
169 <400> SEQUENCE: 2	
170 Met Val Val Gly Met Thr Met Ala Gly Asp Gln Glu Leu Glu Leu Arg	
171 1 5 10 15	
173 Phe Asp Val Pro Leu Tyr Thr Leu Ala Glu Ala Ser Arg Tyr Leu Val	
174 20 25 30	
176 Val Pro Arg Ala Thr Leu Ala Thr Trp Ala Asp Gly Tyr Glu Arg Arg	
177 35 40 45	
179 Pro Ala Asn Ala Pro Ala Val Gln Gly Gln Pro Ile Ala Phe Asp Ala	
180 50 55 60	
182 Tyr Ser Val Ala Gln Leu Phe Gly Asp Val Thr Gly Ala Arg Val Ala	
183 65 70 75 80	
185 Gly Val Gln Pro Gln Arg His His Ile Arg Pro Val Arg Leu Arg Gly	
186 85 90 95	
188 Pro Leu Gly Gly Val Gly Cys Leu Arg His Pro Arg Gln Phe Ala Gly	
189 100 105 110	
191 Tyr Leu Ser Gln Met Ser Trp Met Thr Ala Thr Arg Arg Leu Arg Asn	
192 115 120 125	
194 Arg His Arg Leu Asp Ser Pro Thr Ala Ser Ser Pro Gly Lys Pro Pro	
195 130 135 140	
197 Ala Leu Thr Pro Ala Thr Asn Pro Met Asn Cys Trp Ile Ser Asp Ser	

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/855,604A

DATE: 07/05/2002  
TIME: 14:47:36

Input Set : D:\37156201.app  
Output Set: N:\CRF3\07052002\I855604A.raw

```

198 145          150          155          160
200 Pro Tyr Ser Arg Ala Val Arg Ala Arg Glu Pro Thr Glu Asp Arg Val
201           165          170          175
203 His Ala Phe Gly Val Asp Arg Thr Ala Pro Gly Val Gly Gly Ala Glu
204           180          185          190
206 Gly Arg Asp Gly Arg Met Thr Asp Arg Arg Gly Arg Glu Leu Pro Gly
207           195          200          205
209 Arg Arg Thr Val Ala Asn Pro Ser Gln Thr Arg Arg Lys Pro Met Lys
210           210          215          220
212 Thr Gly Thr Ala Thr Thr Arg Arg Arg Leu Leu Ala Val Leu Ile Ala
213 225          230          235          240
215 Leu Ala Leu Pro Gly Ala Ala Val Ala Leu Leu Ala Glu Pro Ser Ala
216           245          250          255
218 Thr Gly Ala Ser Asp Pro Cys Ala Ala Ser Glu Val Ala Arg Thr Val
219           260          265          270
221 Gly Ser Val Ala Lys Ser Met Gly Asp Tyr Leu Asp Ser His Pro Glu
222           275          280          285
224 Thr Asn Gln Val Met Thr Ala Val Leu Gln Gln Val Gly Pro Gly
225           290          295          300
227 Ser Val Ala Ser Leu Lys Ala His Phe Glu Ala Asn Pro Lys Val Ala
228 305          310          315          320
230 Ser Asp
233 <210> SEQ ID NO: 3
234 <211> LENGTH: 6
235 <212> TYPE: PRT
236 <213> ORGANISM: Mycobacterium tuberculosis
238 <400> SEQUENCE: 3
239 Met Val Val Gly Met Thr
240   1          5
243 <210> SEQ ID NO: 4
244 <211> LENGTH: 110
245 <212> TYPE: PRT
246 <213> ORGANISM: Mycobacterium tuberculosis
248 <400> SEQUENCE: 4
249 Met Ala Gly Asp Gln Glu Leu Glu Leu Arg Phe Asp Val Pro Leu Tyr
250   1          5          10          15
252 Thr Leu Ala Glu Ala Ser Arg Tyr Leu Val Val Pro Arg Ala Thr Leu
253           20          25          30
255 Ala Thr Trp Ala Asp Gly Tyr Glu Arg Arg Pro Ala Asn Ala Pro Ala
256           35          40          45
258 Val Gln Gly Gln Pro Ile Ala Phe Asp Ala Tyr Ser Val Ala Gln Leu
259           50          55          60
261 Phe Gly Asp Val Thr Gly Ala Arg Val Ala Gly Val Gln Pro Gln Arg
262   65          70          75          80
264 His His Ile Arg Pro Val Arg Leu Arg Gly Pro Leu Gly Gly Val Gly
265           85          90          95
267 Cys Leu Arg His Pro Arg Gln Phe Ala Gly Tyr Leu Ser Gln
268           100         105         110
271 <210> SEQ ID NO: 5

```

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/09/855,604A

DATE: 07/05/2002  
TIME: 14:47:36

Input Set : D:\37156201.app  
Output Set: N:\CRF3\07052002\I855604A.raw

272 <211> LENGTH: 33  
273 <212> TYPE: PRT  
274 <213> ORGANISM: Mycobacterium tuberculosis  
276 <400> SEQUENCE: 5  
277 Met Thr Ala Thr Arg Arg Leu Arg Asn Arg His Arg Leu Asp Ser Pro  
278 1 5 10 15  
280 Thr Ala Ser Ser Pro Gly Lys Pro Pro Ala Leu Thr Pro Ala Thr Asn  
281 20 25 30  
283 Pro  
286 <210> SEQ ID NO: 6  
287 <211> LENGTH: 70  
288 <212> TYPE: PRT  
289 <213> ORGANISM: Mycobacterium tuberculosis  
291 <400> SEQUENCE: 6  
292 Met Asn Cys Trp Ile Ser Asp Ser Pro Tyr Ser Arg Ala Val Arg Ala  
293 1 5 10 15  
295 Arg Glu Pro Thr Glu Asp Arg Val His Ala Phe Gly Val Asp Arg Thr  
296 20 25 30  
298 Ala Pro Gly Val Gly Gly Ala Glu Gly Arg Asp Gly Arg Met Thr Asp  
299 35 40 45  
301 Arg Arg Gly Arg Glu Leu Pro Gly Arg Arg Thr Val Ala Asn Pro Ser  
302 50 55 60  
304 Gln Thr Arg Arg Lys Pro  
305 65 70  
308 <210> SEQ ID NO: 7  
309 <211> LENGTH: 100  
310 <212> TYPE: PRT  
311 <213> ORGANISM: Mycobacterium tuberculosis  
313 <400> SEQUENCE: 7  
314 Met Lys Thr Gly Thr Ala Thr Arg Arg Leu Leu Ala Val Leu  
315 1 5 10 15  
317 Ile Ala Leu Ala Leu Pro Gly Ala Ala Val Ala Leu Ala Glu Pro  
318 20 25 30  
320 Ser Ala Thr Gly Ala Ser Asp Pro Cys Ala Ala Ser Glu Val Ala Arg  
321 35 40 45  
323 Thr Val Gly Ser Val Ala Lys Ser Met Gly Asp Tyr Leu Asp Ser His  
324 50 55 60  
326 Pro Glu Thr Asn Gln Val Met Thr Ala Val Leu Gln Gln Val Gly  
327 65 70 75 80  
329 Pro Gly Ser Val Ala Ser Leu Lys Ala His Phe Glu Ala Asn Pro Lys  
330 85 90 95  
332 Val Ala Ser Asp  
333 100  
336 <210> SEQ ID NO: 8  
337 <211> LENGTH: 961  
338 <212> TYPE: DNA  
339 <213> ORGANISM: Mycobacterium tuberculosis  
341 <220> FEATURE:  
342 <221> NAME/KEY: CDS

RAW SEQUENCE LISTING ERROR SUMMARY                    DATE: 07/05/2002  
PATENT APPLICATION: US/09/855,604A                    TIME: 14:47:37

Input Set : D:\37156201.app  
Output Set: N:\CRF3\07052002\I855604A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:935; Xaa Pos. 3

Invalid Line Length:

The rules require that a line not exceed 72 characters in length. This includes spaces.

Seq#:8; Line(s) 420,424  
Seq#:14; Line(s) 650  
Seq#:169; Line(s) 5660  
Seq#:177; Line(s) 5811  
Seq#:228; Line(s) 7196,7200,7204  
Seq#:238; Line(s) 7407,7411  
Seq#:290; Line(s) 8869  
Seq#:297; Line(s) 9097,9117,9125,9141  
Seq#:310; Line(s) 9408,9416  
Seq#:347; Line(s) 10644,10648,10656,10660,10664,10668,10672,10676,10680  
Seq#:347; Line(s) 10684,10688,10692,10696,10700,10704,10708,10712,10732  
Seq#:412; Line(s) 13020,13024  
Seq#:463; Line(s) 14779  
Seq#:510; Line(s) 16792,16796,16800  
Seq#:640; Line(s) 22025  
Seq#:702; Line(s) 24419,24431,24435,24439,24443,24455,24459,24463  
Seq#:717; Line(s) 24867  
Seq#:771; Line(s) 26795,26799  
Seq#:784; Line(s) 27150  
Seq#:794; Line(s) 27446,27450  
Seq#:848; Line(s) 29649,29669,29673,29677,29709,29713  
Seq#:864; Line(s) 30193,30245  
Seq#:887; Line(s) 31454,31458,31462

RECEIVED  
JUL 2 4 2002  
OFFICE OF PETITIONS

VERIFICATION SUMMARY  
PATENT APPLICATION: US/09/855,604A

DATE: 07/05/2002  
TIME: 14:47:37

Input Set : D:\37156201.app  
Output Set: N:\CRF3\07052002\I855604A.raw

L:32792 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:935 after pos.:0

RECEIVED  
JUL 24 2002  
OFFICE OF PETITIONS